

FTLA (Fundamental theorem of linear algebra). If A is $n \times n$ then:

- $\mathcal{R}(A) \perp \mathcal{N}(A^H)$ (incredible freebie!)
- The normal equations of Gauss.

$$\begin{bmatrix} I & A \\ A^H & 0 \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix} \quad (\text{Kahan form}),$$

that is

$$r = b - Ax, \quad A^H r = 0,$$

that is

$$A^H A x = A^H b \quad (\text{Gauss form}),$$

are always solvable.

- Every vector b in \mathbb{C}^n can be written uniquely as

$$b = P + r : \quad P \text{ in } \mathcal{R}(A), \quad r \text{ in } \mathcal{N}(A^H).$$

- a) Let $y = Ax$ be in $\mathcal{R}(A)$ and v be in $\mathcal{N}(A^H)$, so $A^H v = 0$. Then, by the reverse order rule for conjugate transposition, $y^H v = (Ax)^H v = x^H A^H v = x^H 0 = 0$. In general, two vectors v, y in \mathbb{C}^n are orthogonal, or perpendicular ($v \perp y$) when $y^H v = 0$.

b) By Gauss, by golly, omitting the permutation matrices P and Q
only for brevity of exposition,

$$A = LU = \begin{bmatrix} L_1 \\ L_2 \end{bmatrix} [U_1 \ U_2] ,$$

$$A^H = U^H L^H, \quad A^H A = U^H L^H L U ,$$

so the Gauss form is

$$U^H L^H L U \underline{x} = U^H L^H \underline{b} .$$

This implies

$$U_1^H L^H L U \underline{x} = U_1^H L^H \underline{b}$$

and, since $U_1^H = \Delta$ is nonsingular,

$$L^H L U \underline{x} = L^H \underline{b} .$$

Now the matrix $L^H L$ is also nonsingular. In fact it is Hermitian positive definite:

$$\underline{x}^H L^H L \underline{x} = (L \underline{x})^H (L \underline{x}) = \|L \underline{x}\|^2 \geq 0 ,$$

with equality iff $L \underline{x} = \underline{0}$, that is, iff $L_1 \underline{x} = \underline{0}$ and $L_2 \underline{x} = \underline{0}$, that is, since L_1 is nonsingular, iff $\underline{x} = \underline{0}$.

So $L^H L \underline{c} = L^H \underline{b}$ is uniquely solvable for \underline{c} , and then

$$U \underline{x} = [U_1 \ U_2] \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix}$$

$$= U_1 \underline{x}_1 + U_2 \underline{x}_2 = \underline{c}$$

is always solvable for \underline{x} . The steps are reversible. This "algorithm" even gives the general solution \underline{x} of Gauss' normal equations. This gives the main part of the FTLA, in theory. There are variants of the algorithm that are "more stable numerically".

c) The existence of $\underline{P} := A\underline{x}$ and \underline{r} are solved by b). To show the uniqueness assume a second such decomposition

$$\underline{b} = \underline{P}_0 + \underline{r}_0, \quad \underline{P}_0 \text{ in } R(A), \quad \underline{r}_0 \text{ in } \gamma(A^H).$$

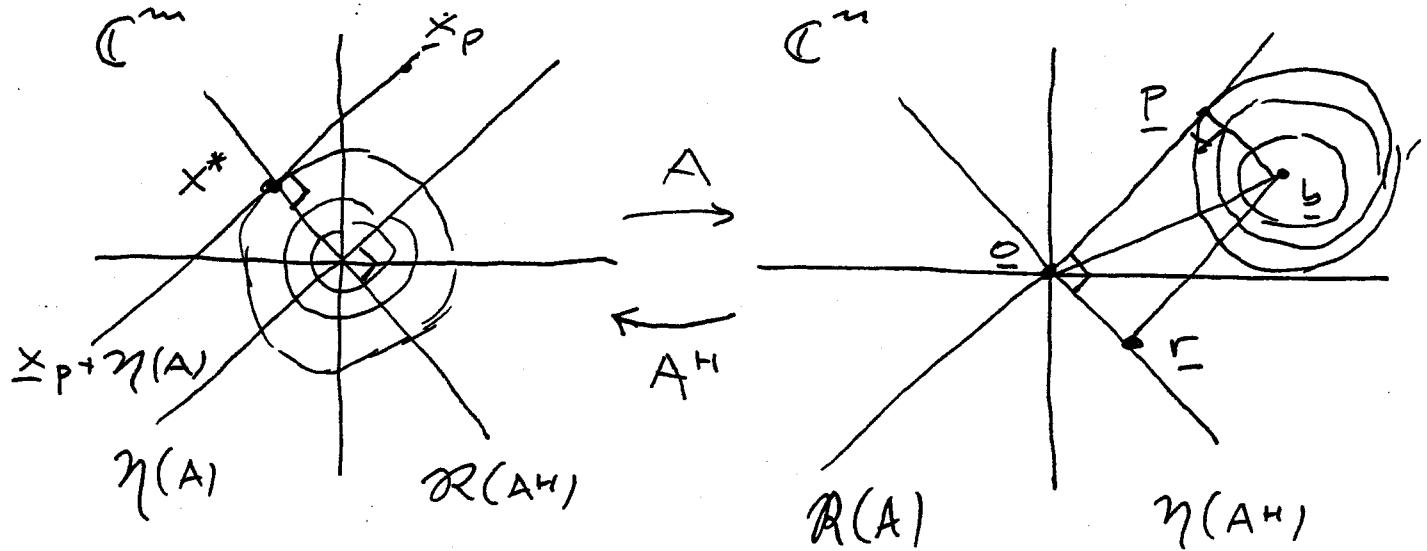
Then

$$\underline{P} - \underline{P}_0 = \underline{r}_0 - \underline{r}. \quad (*)$$

Now, by the linearity of $A\underline{x}$ in \underline{x} , the subsets $R(A)$ and $\gamma(A^H)$ of \mathbb{C}^n are subspaces of \mathbb{C}^n , that is they are closed under the formation of linear combinations. So the $(*)$ vector is in both $R(A)$ and $\gamma(A^H)$ and is thus orthogonal with

itself, by a). So $\underline{P}_0 = \underline{P}$, $\underline{r}_0 = \underline{r}$. ■

The BIG PICTURE:



x^* = shortest solution of
least squares problem
 $\|\underline{b} - A \underline{x}\|^2$ = minimum,
minimizes $\|\underline{x}\|$ over all
such solutions

$\underline{P} = A \underline{x}_p$, a particular \underline{x}_p

= orthogonal projection of
 \underline{b} onto $R(A)$

$\underline{r} = \underline{b} - \underline{P}$ = residual vector

$\underline{x}^* = A^I \underline{b}$, $A^I = \underline{\text{Moore-Penrose}}$
pseudoinverse of A.

Again, there are "formulas" for A^I ,
but if A^{-1} is "bad" then A^I can't be better!

ftla (little FTLA). $\underline{A} = \underline{a} \neq \underline{0}$, $m = 1$.

Least squares problem. Find the closest point \underline{p} in the line $R(\underline{a})$ through $\underline{0}$ in \mathbb{C}^n to a given point \underline{b} in \mathbb{C}^n . That is, choose \underline{x} in \mathbb{C} to minimize the sum of squared absolute values

$$f(\underline{x}) = \|\underline{r}\|^2 = \|\underline{b} - \underline{a}\underline{x}\|^2.$$

Complete the square:

$$\begin{aligned} f(\underline{x}) &= (\underline{b} - \underline{a}\underline{x})^H(\underline{b} - \underline{a}\underline{x}) \\ &= (\underline{b}^H - \bar{\underline{x}}\underline{a}^H)(\underline{b} - \underline{a}\underline{x}) \\ &= \underline{b}^H\underline{b} - \underline{b}^H\underline{a}\underline{x} - \bar{\underline{x}}\underline{a}^H\underline{b} + \bar{\underline{x}}\underline{a}^H\underline{a}\underline{x} \\ &= \|\underline{b}\|^2 - 2\operatorname{Re}\underline{a}^H\underline{b}\bar{\underline{x}} + \|\underline{a}\|^2|\underline{x}|^2 \\ &= \|\underline{a}\|^2\left(|\underline{x}|^2 - 2\operatorname{Re}\frac{\underline{a}^H\underline{b}}{\|\underline{a}\|^2}\bar{\underline{x}} + \frac{\|\underline{b}\|^2}{\|\underline{a}\|^2}\right). \end{aligned}$$

Now, for complex c , and complex x ,

$$\begin{aligned} |\underline{x} - c|^2 &= (\bar{\underline{x}} - \bar{c})(\underline{x} - c) \\ &= |\underline{x}|^2 - \bar{\underline{x}}c - \bar{c}\underline{x} + |c|^2 \\ &= |\underline{x}|^2 - 2\operatorname{Re}c\bar{\underline{x}} + |c|^2. \end{aligned}$$

Choose

$$c = \frac{\underline{a}^H \underline{b}}{\underline{a}^H \underline{a}}$$

and work backwards:

$$\begin{aligned} f(x) &= \|\underline{a}\|^2 \left(\left| x - \frac{\underline{a}^H \underline{b}}{\underline{a}^H \underline{a}} \right|^2 + \frac{\|\underline{b}\|^2}{\|\underline{a}\|^2} - \frac{|\underline{a}^H \underline{b}|^2}{\|\underline{a}\|^4} \right) \\ &= \|\underline{a}\|^2 \left| x - \frac{\underline{a}^H \underline{b}}{\underline{a}^H \underline{a}} \right|^2 + \frac{\|\underline{a}\|^2 \|\underline{b}\|^2 - |\underline{a}^H \underline{b}|^2}{\|\underline{a}\|^2} \\ &\geq \frac{\|\underline{a}\|^2 \|\underline{b}\|^2 - |\underline{a}^H \underline{b}|^2}{\|\underline{a}\|^2}, \end{aligned}$$

with equality iff

$$x = \frac{\underline{a}^H \underline{b}}{\underline{a}^H \underline{a}}.$$

The closest point in $\mathcal{R}(\underline{a})$ to \underline{b} is

$$P = \underline{a} \underline{x} = \frac{\underline{a} \underline{a}^H}{\underline{a}^H \underline{a}} \underline{b} =: P_{\underline{a}} \underline{b},$$

no matter what the vector \underline{b} is in \mathbb{C}^n . The matrix $P_{\underline{a}}$ projects each given \underline{b} onto its own \underline{p} in $\mathcal{R}(\underline{a})$.

We have

$$P_{\underline{a}} = P_{\underline{a}}^H, \quad P_{\underline{a}}^2 = P_{\underline{a}}.$$

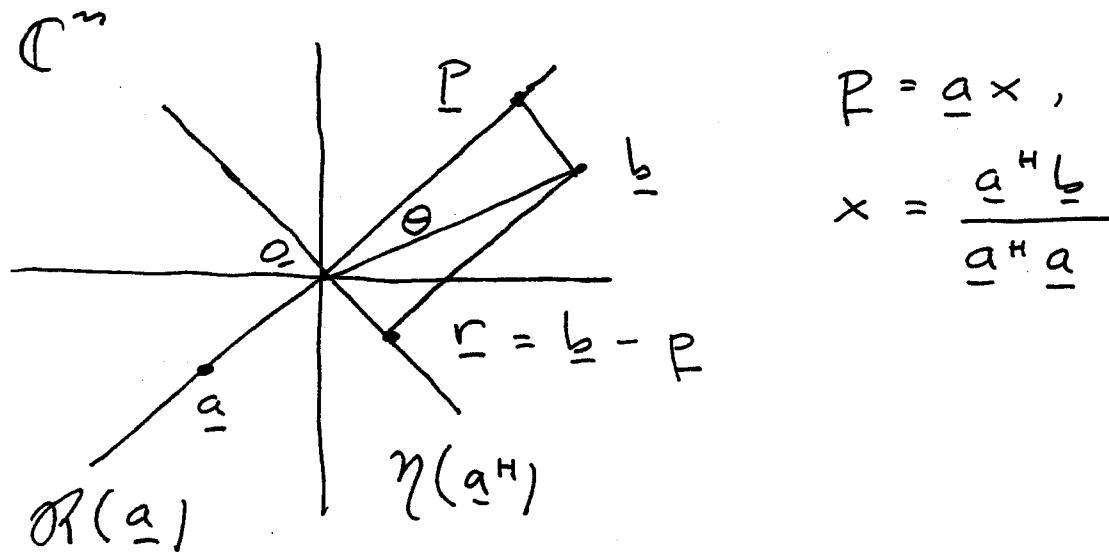
Now always $f(x) \geq 0$, and the minimum value is zero iff \underline{b} is already in $R(\underline{a})$. This gives

Cauchy's inequality. For any vectors \underline{a} and \underline{b} in \mathbb{C}^n ,

$$|\underline{a}^H \underline{b}| \leq \|\underline{a}\| \|\underline{b}\|,$$

with equality if and only if one of \underline{a} or \underline{b} is a scalar multiple of the other.

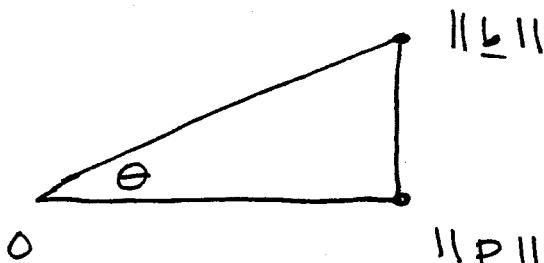
The little picture is



$$\begin{aligned} \underline{P} &= \underline{a}x, \\ x &= \frac{\underline{a}^H \underline{b}}{\underline{a}^H \underline{a}} \end{aligned}$$

Let θ be the acute ($\leq 90^\circ$) angle between the lines determined by \underline{a} and \underline{b} , assuming also $\underline{b} \neq 0$ now. The three vectors \underline{a} , $\underline{P} = \underline{a}x$ and \underline{b}

determine a two dimensional plane in \mathbb{C}^n (provided b is not in $R(a)$). Extract the key 2 -real D triangle:



$$\|P\| = \|a\| \times 1 = \frac{|a^H b|}{\|a\|}$$

$$= \|b\| \cos \theta$$

Hence, the angle θ between the lines $R(a)$ and $R(b)$ is given by

$$\cos \theta = \frac{|a^H b|}{\|a\| \|b\|} : 0 \leq \theta \leq \frac{\pi}{2}.$$

a and b are orthogonal, or perpendicular ($a \perp b \iff b \perp a$) if $\theta = \frac{\pi}{2}$, that is if $a^H b = 0$ ($\iff b^H a = 0$).

The normal equation just says that r and a must be orthogonal. Similarly in the general case $n \geq 1$. Finally, orthogonality gives the

Pythagorean theorem:

$$\|b\|^2 = \|P\|^2 + \|r\|^2.$$